

Archivierung mediengebundener Formate: Audio-CDs in Dateien umwandeln

Objekt der digitalen Archivierung sind Dateien. Audio-CDs (wie auch andere Datenträger aus dem audiovisuellen Bereich, z.B. Digibeta-Kassetten) speichern ihre Daten jedoch in einem medientypischen Datenstrom. Zur Archivierung muss dieser Datenstrom mit einem Ripper-Programm auf archivisch akzeptable Weise in Dateien umgewandelt werden.

1. Das Problem: Datenstrom statt Datei	1
2. Die Lösung: CD-Ripper	1
3. Was zu beachten ist	1
4. Demonstration: Exact Audio Copy	2
5. Ressourcenfragen	3
6. Weiterführende Links	3

1. Das Problem: Datenstrom statt Datei

Eine Audio-CD enthält keine Dateien, sondern einen zusammenhängenden Datenstrom, welcher vom Abspielgerät bzw. der Abspielsoftware interpretiert wird *i*. Man spricht von einem mediengebundenen Format. Objekte der digitalen Archivierung sind jedoch Dateien. Dies gilt für den Text- und Bildbereich sowie für strukturierte Daten unhinterfragt. Auch der Bereich der Audio-Archivierung ist inzwischen vollständig dateibasiert; im Videobereich ist die Umstellung im Gang (siehe Felix Rauh, Archivierung mediengebundener Formate: Audio-CDs, Digibeta. KOST-Werkstattbericht, 17. November 2008). Im KOST-Projekt *Sauver les CD* ist dieser Grundsatz für die Langzeitarchivierung von Daten auf ephemeren digitalen Datenträgern ausdrücklich bestätigt worden (siehe dazu die Protokolle der Workshops vom 09.06.2008 und 25.09.2008). Aus dem an die CD gebundenen Datenstrom müssen zur Archivierung also einzelne Dateien erstellt werden *ii*.

i Siehe dazu Wikipedia: [Compact Disc Digital Audio](#). Wenn eine Audio-CD in das CD-Laufwerk eines PCs eingelegt wird, listet MS Windows ihren Inhalt als einzelne Dateien vom Typ `.cda` auf. Tatsächlich handelt es sich dabei aber einzig um Indexdateien, die auf die entsprechende Stelle auf der CD verweisen und ohne diese nutzlos sind (siehe Wikipedia: [Compact Disc Audio Track](#)). Unter MacOS werden Dateien zunächst ebenfalls nur vorgespiegelt; wenn diese kopiert werden, werden sie jedoch im Hintergrund gerippt, d.h. der entsprechende Datenstrom wird in eine Datei umgewandelt.

ii Alternativ besteht die Möglichkeit, die CD als Ganzes als Datei zu archivieren. Dies wird erreicht mit einem sog. CD-Image, d.h. einem Speicherabbild der CD im Format ISO 9660 (siehe Wikipedia: [ISO-Abbild](#)). Dieses ist nicht direkt lesbar, sondern muss zunächst wieder auf eine CD geschrieben bzw. in einem virtuellen CD-Laufwerk gemountet werden. Weil eine CD im archivischen Kontext häufig eine zufällige Einheit oder ein reines Transportmedium ist, und weil CD-Images schlechter zugänglich sind als Dateien, sollte bei der Archivierung auf diese Variante verzichtet werden.

2. Die Lösung: CD-Ripper

Für die Umwandlung eines Audiodatenstroms in eine Audiofile wird ein Programm vom Typ *CD-Ripper* benötigt. Die grosse Verbreitung digitaler portabler Musikplayer hat das CD-Rippen zu einer beinahe alltäglichen Beschäftigung gemacht. Entsprechend sind unzählige *CD-Ripper* verfügbar (siehe dazu die Liste in Wikipedia: [CD Ripper](#)). Viele davon sind in Betriebssystemen bzw. Audiosoftware integriert oder als Freeware- bzw. Open-Source-Programme erhältlich.

3. Was zu beachten ist

Bei der Wahl eines Produkts für die archivische Anwendung aus der grossen Auswahl an CD-Rip-Software sind die folgenden Punkte zu beachten:

- Als Archivformat für Audiodateien wird zurzeit WAVE empfohlen *iii*. Der CD-Ripper muss also dieses Format ausgeben können. Diese Anforderung ist trivial, da praktisch sämtliche verfügbaren Ripper WAVE-Dateien erzeugen können.
- Um eine sichere und korrekte Datenübernahme von der CD in die Audiodateien gewährleisten zu können, muss der CD-Ripper einerseits Lesefehler vermeiden, andererseits mit beschädigten Daten umgehen können. Lesefehler werden vermieden durch verschiedene Verifikationsmechanismen. Wenn beschädigte Daten mit der Fehlerkorrektur des Laufwerks oder der Software repariert werden können, muss der Fehler genau dokumentiert werden (siehe dazu Wikipedia, [Accurate/secure rip](#)).
- Der Rip-Vorgang muss in einem aussagekräftigen Report dokumentiert werden. Dabei sind insbesondere

iii Siehe den Katalog archivischer Dateiformate (KaD) der KOST, <http://www.kost-ceco.ch/wiki/whelp/KaD/>, sowie andere ähnliche Kataloge in der dortigen Bibliografie.

unlesbare Stellen und Lesefehler genau zu bezeichnen.

Es existieren mehrere kommerzielle und Freeware-Produkte, die diese Anforderungen erfüllen. Eine entsprechende Liste findet sich unter Wikipedia, [Accurate/secure rip](#). Die KOST hat nicht alle diese Ripper getestet und macht keine Aussage über ihre Eignung. Im Folgenden präsentieren wir eine mögliche Wahl: *Exact Audio Copy*.

4. Demonstration: Exact Audio Copy

Exact Audio Copy (EAC) ist Freeware und kann von <http://www.exactaudiocopy.de/> heruntergeladen werden. (Das Programm liegt zurzeit in V0.99 prebeta 4 vor.) Die .exe-Datei wird nach dem Download mittels Doppelklick installiert. Beim Start von EAC soll sich eine CD im optischen Laufwerk des Computers befinden.

Beim erstmaligen Ausführen von EAC auf einem Computer unterstützt der *Setup Wizard* die Konfiguration. Dabei werden die folgenden Schritte ausgeführt:

- Laufwerksauswahl: Der Wizard erkennt die vorhandenen optischen Laufwerke; es muss manuell ausgewählt werden, welche davon in die Konfiguration einzubeziehen sind.
- Priorität: Es kann zwischen Präzision (*accurate results*) und Geschwindigkeit (*speed*) gewählt werden; empfohlen wird ersteres.
- Laufwerksinformationen: Zur Konfiguration können die Informationen über das entsprechende Laufwerk aus der EAC-Datenbank verwendet oder die Laufwerke analysiert werden.
- Identifikation des besten Laufwerks (falls der Computer mehrere Laufwerke umfasst).
- Wahl eines Kompressionsformats: ist für archivische Anwendung in der Regel irrelevant.
- Festlegung des Namensschemas.
- Modusauswahl: Der *Beginner Mode* blendet einige Optionen aus; es wird empfohlen, direkt den *Expert Mode* zu verwenden.

Damit ist die Konfiguration abgeschlossen. Über die Menüs können nun weitere Optionen gewählt werden:

- EAC > EAC Options > Extraction: Die *Error recovery quality* bestimmt die Anzahl Leseversuche im *Secure Mode* (siehe unten). Sie sollte auf *High* gesetzt werden.
- EAC > Drive Options > Extraction Method: Es besteht die Wahl zwischen

Secure Mode, *Synchronised Mode* und *Burst Mode*. Im *Secure Mode* wird jeder Sektor zunächst zweimal gelesen.

Wenn die Werte nicht übereinstimmen, werden 16 weitere Lesevorgänge durchgeführt. Der Sektor gilt als richtig gelesen, wenn mindestens 8 dieser 16 Resultate übereinstimmen. Je nach Einstellungen (siehe oben) wird diese Serie von 16 Lesevorgängen bei Bedarf bis zu fünfmal wiederholt. Es wird empfohlen, die Verwendung der *C2 error information* nur zu aktivieren, falls beim Rippen die Methode *Test & Copy* gewählt wird (siehe unten), da diese Fehlerinformationen von gewissen Laufwerken nicht immer zuverlässig angegeben werden. *Synchronised* und *Burst Mode* haben weniger bis keine Fehlerkontrollmechanismen.

- EAC > Drive Options > Drive: Schaltfläche *Autodetect read command now* betätigen.
- EAC > Drive Options > Gap Detection: Hier besteht Experimentationsspielraum. Empfehlenswert ist die Verwendung der sichersten Methode, wenn dies die Analyse nicht zu sehr verlangsamt.
- EAC > Drive Options > Offset/Speed: Hier kann die Lesegeschwindigkeit eingestellt werden. Diese Einstellung hat allerdings wenig Auswirkung auf die tatsächliche Geschwindigkeit, welche vom Programm automatisch angepasst wird.
- EAC > Profiles erlaubt es, Einstellungen als Profil abzuspeichern oder zu übernehmen.
- Compression Options ermöglichen die Konfiguration der Datenkomprimierung.

Zum Rippen einer CD werden sämtliche Tracks der CD markiert und dann im Menü *Action* entweder *Copy Selected Tracks* oder *Test & Copy Selected Tracks* gewählt. Bei der zweiten Variante wird jeder Track zunächst testweise und dann richtig ausgelesen. Wenn die Checksummen beider Lesevorgänge identisch sind, kann die Extrahierung als korrekt gelten. Am Schluss erscheint ein *Status Report*, der mit den gerippten Dateien zusammen gespeichert werden soll. Er detailliert die Resultate des Auslesevorgangs. Zu beachten ist, dass eine *Track quality* von unter 100% nur bedeutet, dass gewisse Sektoren mehrfach gelesen werden mussten. Solange keine "Suspicious positions" aufgeführt sind, wurden sämtliche allfälligen Fehler korrigiert.

Anmerkung: Der Hauptanwendungsbereich von EAC ist das Rippen von Musik-CDs aus Massenproduktion, währenddem im Archivbereich in der Regel Unikate vorliegen. Deshalb machen zwei Optionen der Software für Archive keinen Sinn und sollten deaktiviert werden:

- *Freedb*: Verschiedene öffentlich zugängliche Datenbanken ermöglichen es, Metadaten (Interpret, Titel, etc.) zu gerippten CDs direkt zu importieren.
- *AccurateRip*: EAC verwendet zur Verifizierung der Rip-Resultate das Plugin *AccurateRip* (Siehe <http://www.exactaudiocopy.de/en/index.php/overview/basic-technology/accurate-rip/> sowie <http://www accuraterip.com/>). Dieses vergleicht die Checksumme einer gerippten Datei mit den Checksummen der gleichen Datei (d.h. des gleichen *Audiotracks*), welche andere User gerippt haben. So wird das von Ihnen erreichte Resultat erhärtet.

Für weitergehende Informationen verweisen wir auf die EAC-Website, insbesondere auf die [Dokumentation](#). Eine illustrierte Bedienungsanleitung mit ausführlichen Erklärungen zu *Exact Audio Copy* bietet Frank Bicking, [Audio-CDs rippen mit Exact Audio Copy](#).

5. Ressourcenfragen

Audiodateien im unkomprimierten WAVE-Format sind relativ voluminös. Als Faustregel kann gelten: 1 Minute Audio entspricht 10 MB Speicherplatz. Durch Datenkompression kann dieser Umfang je nach Kompressionsrate auf rund 10-20% verkleinert werden. Allerdings gelten in der Langzeitar Archivierung unkomprimierte Formate als *best practice*.

Die Dauer des Rip-Vorgangs ist abhängig von den Einstellungen und dementsprechend von den Anforderungen an die Sicherheit. Grundsätzlich erhöht der *Secure Mode* die notwendige Zeit auf über das Doppelte, weil die CD nicht kontinuierlich, sondern jeder Sektor zweimal gelesen wird, und die Verwendung von *Test & Copy* ver-

doppelt die Zeit nochmals, weil der gesamte Lesevorgang zweimal durchgeführt wird. Weitere Einstellungen verlängern oder verkürzen den Rip-Vorgang entsprechend. Tests mit einer problemlos lesbaren CD ergaben, dass das schnellstmögliche Rippen im *Burst-Mode* ungefähr neunfache Geschwindigkeit erreichte, währenddem der sicherstmögliche Modus nur auf knapp zweifache Geschwindigkeit kam. Kratzer oder andere Defekte verlängern die Dauer des Rip-Vorgangs erheblich. Als guter Kompromiss kann die Verwendung des *Burst Mode* zusammen mit *Test & Copy* gelten. Bei nicht beschädigten oder übermäßig alten CDs ist der *Burst Mode* zuverlässig genug, und wenn die durch *Test & Copy* erhaltenen Checksummen identisch sind, ist eine hinreichende Sicherheit erreicht. Bei unterschiedlichen Checksummen muss die CD erneut gerippt werden, besser dann im *Secure Mode*.

Größere Bestände werden idealerweise in Stapelverarbeitung gerippt. Dazu wird ein CD-Wechsler benötigt, welcher über die Kommandozeile angesprochen werden kann (Informationen dazu bei der KOST-Geschäftsstelle). Alternativ empfiehlt es sich, das Rippen durch einen externen Dienstleister nach den Vorgaben des Archivs durchführen zu lassen.

6. Weiterführende Links

- Frank Bicking, Audio-CDs rippen mit Exact Audio Copy. AudioHQ 2003-2008, <http://www.audiohq.de/index.php?showtopic=47>
- The Essential Ripping Guide for EAC <http://users.fulladsl.be/spb2267/>
- Memoriav (ed.), Ton. Die Erhaltung von Tondokumenten. Memoriav Empfehlungen. 2008. http://de.memoriav.ch/dokument/Empfehlungen/empfehlungen_ton_de.pdf
- Teqnilogik.com
Exact Audio Copy Guide
Version 2.0, June 2006
<http://www.teqnilogik.com/tutorials/eac.shtml>